

1 **Title: Dissecting Biological Pathways of Psychopathology using Cognitive Genomics**

2
3 **Authors:** Max Lam^{1,2,3,4,5}, Chia-Yen Chen⁶, W. David Hill⁷, Charley Xia⁷, Ruoyu Tian⁶, Daniel F.
4 Levey^{8,9}, Joel Gelernter^{9,10}, Murray B. Stein^{11,12,13}, Biogen Biobank team⁶, Alexander S.
5 Hatoum^{14,15,16}, Hailiang Huang^{3,4}, Anil K. Malhotra^{1,2,17}, Heiko Runz⁶, Tian Ge^{3,18,19}, Todd
6 Lencz^{1,2,17}

7
8 **Affiliations:**

- 9 1. Division of Psychiatry Research, The Zucker Hillside Hospital, Northwell, Glen Oaks, NY, USA
- 10 2. Institute of Behavioral Science, Feinstein Institutes for Medical Research, Manhasset, NY, USA
- 11 3. Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA
- 12 4. Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA
- 13 5. Institute of Mental Health, Singapore, Singapore
- 14 6. Translational Biology, Research and Development, Biogen Inc., Cambridge, MA, USA
- 15 7. Lothian Birth Cohorts group, Department of Psychology, University of Edinburgh, Edinburgh, UK
- 16 8. Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA
- 17 9. VA Connecticut Healthcare Center, West Haven, CT, USA
- 18 10. Departments of Psychiatry, Genetics, and Neuroscience, Yale University School of Medicine, New Haven, CT,
19 USA
- 20 11. VA San Diego Healthcare System, San Diego, CA, USA
- 21 12. Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA
- 22 13. Herbert Wertheim School of Public Health and Human Longevity Science, University of California San Diego,
23 La Jolla, CA, USA
- 24 14. Institute for Behavioral Genetics, University of Colorado Boulder
- 25 15. Department of Psychology and Neuroscience, University of Colorado Boulder
- 26 16. Department of Psychiatry, University of Washington St. Louis Medical School
- 27 17. Departments of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Norwell,
28 Hempstead, NY, USA
- 29 18. Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General
30 Hospital, Boston, MA, USA
- 31 19. Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

32
33 *To whom correspondence should be addressed:

34 Todd Lencz, Zucker Hillside Hospital, Division of Psychiatry Research, 75-59 263rd Street, Glen
35 Oaks, NY, 11004, USA, E-mail: tlencz@northwell.edu.

36
37
38

1
2
3
4
5
6
7
8
9
10
11
12
13
14

Abstract

Cognitive deficits are known to be related to most forms of psychopathology. Here, we perform local genetic correlation analysis as a means of identifying independent segments of the genome that show biologically interpretable pleiotropic associations between cognitive dimensions and psychopathology. We identified collective segments of the genome, which we call “meta-loci”, that showed differential pleiotropic patterns for psychopathology relative to either *General Cognitive Ability* (GCA) or *Non-Cognitive Skills* (NCS). We observed that neurodevelopmental gene sets expressed during the prenatal-early childhood predominated in GCA-relevant meta-loci, while post-natal synaptic gene sets were more involved in NCS-relevant meta-loci. Notably, we found that GABA-ergic, cholinergic, and glutamatergic genes drove pleiotropic relationships within dissociable NCS meta-loci.

1 Cognitive impairment is one of the core features of psychopathology and is associated
2 with the debilitating nature of many psychiatric disorders^{1,2}. In schizophrenia, for example,
3 cognitive impairments are predictive of known functional impairments even in the prodromal
4 stage of the illness^{3,4}. Cognitive deficits are not only confined^{3,4} to adult psychiatric illnesses, but
5 also extend to childhood disorders like Autism Spectrum Disorders (ASD) and Attention-Deficit
6 Hyperactivity Disorder (ADHD)⁵. Individuals who suffer from psychiatric disorders tend to report
7 sequelae of cognitive problems throughout their lifetime⁶. In many cases, there is an emergence
8 of cognitive deficits before a formal diagnosis of mental illness⁷.

9 Prior to the era of well-powered GWAS in psychiatry, the idea of using cognitive function
10 as an endophenotype to understand the biology of psychopathology was proposed⁸. We
11 presented the first molecular genetic evidence for overlap between general cognitive ability and
12 schizophrenia⁹. Since then, evidence suggesting widespread pleiotropy across
13 psychopathologic traits has emerged, indicating shared biological mechanisms^{10,11}. Pleiotropy, a
14 phenomenon where a genetic variant might affect several traits at once, appears to be
15 ubiquitous in biology; 44% of the loci reported within the GWAS catalog have been shown to be
16 associated with more than one trait¹² (although in some cases this may be a function of linkage
17 disequilibrium rather than true pleiotropy¹³). A more recent study indicated that trait associated
18 loci cover more than half of the genome, among which 90% implicate multiple traits¹⁴.

19 We recently exploited pleiotropy to dissect biology underlying the counterintuitive
20 positive genetic correlation between educational attainment and schizophrenia, and were able
21 to parse separate neurodevelopmental and synaptic mechanisms underlying the disorder,
22 based on association patterns within GWAS significant loci¹⁵. Shortly after, Demange and
23 colleagues¹⁶ demonstrated that it was possible to leverage global genetic correlation within a
24 structural equation modelling framework to derive a latent *Non-Cognitive Skills* (NCS) factor by
25 removing variance related to *General Cognitive Ability* (GCA) from educational attainment
26 GWAS. The ensuing NCS factor showed positive genetic correlation with schizophrenia,

1 consistent with our earlier findings. These two reports, utilizing somewhat orthogonal but
2 complementary approaches, leveraged pleiotropic phenomena to study the intersection of
3 cognition and psychopathology. Nevertheless, these studies are limited by following a global
4 genetic correlation approach¹⁷ on the one hand, and a SNP-by-SNP approach^{15,16} on the other.
5 As has been demonstrated¹⁸, the assumption that genetic correlations for complex traits are
6 homogeneously distributed across independent genomic regions may not be true. At the same
7 time, typical locus-based GWAS comparisons tend to be defined by "top" SNPs followed by LD
8 clumping; this invariably results in loss of information from regions of the genome that fall short
9 of genome-wide significance. Studies using partitioned heritability or gene set analysis have
10 demonstrated the biological relevance of regions of the genome that do not contain genome-
11 wide significant loci¹⁹. In the present study, we develop a novel method, intermediate to global
12 and SNP-based approaches, that examines the structure of local genetic correlations across the
13 genome and identifies "meta-loci", which we define as combined genomic segments sharing
14 pleiotropic patterns. We then interrogate these meta-loci, using rich annotations and gene set
15 analysis, to identify and dissociate biological pathways underlying different patterns of cognitive-
16 psychopathologic pleiotropy.

17 **Results**

18 *Study Design and Methods Overview*

19 We have recently reported the largest GWAS meta-analysis for GCA (N=373,617) and
20 utilized this well-powered phenotype for pleiotropic analysis with an expanded list of
21 psychopathology phenotypes²⁰. Expanding on earlier analytic strategies^{15,16}, we carried out
22 three broad stages of analyses (Fig-1).

23 First, as both a data reduction step and a benchmark for subsequent steps, we
24 examined the global pleiotropic relationships between multiple cognitive and psychopathological
25 phenotypes via LD score regression^{21,22}. We amalgamated all available summary statistics from

1 recent GWAS of cognitive function (N = 19). These included GCA²⁰, and other reported
2 derivatives of cognition including NCS¹⁶, Executive Function²³, and GWAS of individual
3 cognitive tests administered as part of the UK Biobank. Similarly, traits related to psychiatric
4 illnesses or psychopathology in the recent studies from the Psychiatric Genomics Consortium,
5 UK Biobank, Million Veterans Program and elsewhere were curated (N = 17; Table-S1). Data
6 reduction approaches indicated that it was appropriate to focus on GCA and NCS traits for the
7 purpose of the current study (Fig-S1). We re-estimated NCS using the largest GCA GWAS²⁰ to
8 increase statistical power (Supplementary Information).

9 Second, we carried out local genetic correlation analyses between each of the
10 psychopathology traits and GCA, and separately between each psychopathology trait and NCS.
11 A series of positive and negative local genetic correlation patterns emerged across these
12 analyses, which we then classified via the “meta-locus” approach. A meta-locus was defined as
13 a set of LD-independent regions that showed similar local genetic correlation profiles across
14 psychopathological traits across the genome. There were 15 distinct meta-loci (Median Length
15 = 34.24 Mb) identified for GCA and 12 identified for NCS (Median Length = 23.19 Mb).

16 Third, GCA and NCS GWAS summary statistics were functionally prioritized to identify
17 genes and biological mechanisms harbored within the meta-loci. A series of Gene-Based
18 Genome-Wide Association (GBGWA) and Transcriptome-Wide Association (TWA) strategies
19 were applied to GCA and NCS GWAS summary statistics. We leveraged brain eQTLs from a
20 range of databases, including GTEx v8 brain tissue expression²⁴, Brain-eQTL-meta²⁵, and
21 PsychEncode^{26,27} eQTL databases that indexes brain (Online Methodology and Supplementary
22 Information). We adopted a broadly inclusive approach to the gene prioritization stage of the
23 analysis, excluding from consideration only those genes with no supporting evidence from any
24 of these procedures. Next, we performed a series of gene set analyses²⁸⁻³¹ and gene scoring
25 procedures³², to which we applied strict filtering criteria in order to arrive at a high-confidence
26 biological characterization of each meta-locus. We evaluated the high-confidence genes

1 emanating from the GBGWA, TWA and gene set analysis approaches for longitudinal gene
2 expression across lifespan to further differentiate neurodevelopmental vs adult synaptic
3 mechanisms. Finally, we annotated driver genes with information on the propensity for
4 psychiatric or nootropic drug re-purposing (Fig-1, Online Methodology).

5 **Stage 1: Global genetic architectures of cognitive and psychopathological traits**

6 A wrapper function within GenomicSEM¹¹ was used to conduct LD score regression^{21,22}
7 and create a global 19 x 17 genetic correlation matrix of cognitive and psychopathological traits
8 (Table-S2). Two-dimension reduction techniques, PCA and partitioned cluster analyses (K-
9 Medoid) dimension reduction techniques were applied to the global genetic correlation matrix to
10 identify underlying association patterns. The PCA resulted in a “Cognitive Map” (Fig-S1a). The
11 first principal-component captured the similarity of each cognitive trait to general cognitive ability
12 (GCA) in context of its relationship to the vector of 19 psychopathological traits (Fig-S1b). The
13 second principal-component represented the degree to which a cognitive trait is similar to NCS
14 given its relationship to the vector of 19 psychopathological traits (Fig-S1c). Consequently, for
15 all subsequent analyses, we focused on GCA and NCS as the primary cognitive phenotypes of
16 interest; we leave a more detailed exploration of the cognitive phenotypic space to future work.
17 We re-extracted the NCS latent factor by using GWAS-by-subtraction parameters¹⁶, and a
18 better powered GCA meta-analysis²⁰. We confirmed that our newly calculated GCA and NCS
19 factors were globally similar to those originally reported by Demange et al.¹⁶ ($r_g=1$ using LD
20 score regression). We utilized the current versions of GCA and NCS in subsequent sections
21 detailing functional annotation, gene prioritization and gene set analyses. At the global level, we
22 also observed that psychopathological traits separated into five best fitting clusters (Table-S2c)
23 that varied according to the degree of the strength of relationship with GCA and NCS (Fig-S1d,
24 inset). There was much more variation across psychopathological traits than the cognitive traits,
25 so we decided to analyze each trait separately in the local genetic correlation analyses
26 described below. However, we utilized the clustering results depicted in Fig-S1d as background

1 information for our interpretations of subsequent results (and as color-coding in subsequent
2 figures in this report).

3 **Stage 2: Investigating pleiotropic relationships between cognitive dimensions and** 4 **psychopathology**

5 Local genetic correlation analyses were carried out across 2353 LD-independent regions
6 of the genome using ρ -HESS¹⁸. 4,469,149 SNPs were included for GCA and 4,372,398 SNPs
7 were included for NCS. Region specific heritability estimates of GCA and NCS were expectedly
8 small (median $h^2_{\text{GCA-Region}} = 8.36\text{e-}5$, median $h^2_{\text{NCS-Region}} = 1.18\text{e-}4$, Table-S3). The sum of
9 heritability across LD independent regions was consistent with previous reports for these
10 phenotypes ($h^2_{\text{GCA}} = 0.23$, $h^2_{\text{NCS}} = 0.31$) (Supplementary Information).

11 We observed widespread pleiotropy for both GCA and NCS with psychopathology
12 (Table-S4/S5). In general, global estimates of genetic correlations (via LD-score regression)
13 were highly comparable to summed local estimates derived from ρ -HESS, except in the case of
14 Anorexia Nervosa and OCD, for which ρ -HESS recovered a greater degree of correlation with
15 NCS (Fig-S2). Aside from re-capitulating global R_g trends across psychopathology phenotypes,
16 several noteworthy observations emerge from the local genetic correlations. First, for each
17 psychopathological trait, local genetic correlations were not in the expected direction, across
18 LD-independent regions, indicated by global genetic correlations (Fig-S3a/b). Second, the
19 magnitude of the local genetic correlation signals is markedly stronger for schizophrenia than for
20 any other trait. Third, schizophrenia is the only trait (potential exception of Anxiety_MVP) that
21 demonstrates strong, widespread differences in direction between GCA and NCS correlations
22 (Fig-S2, Table-S6). Fourth, local genetic correlations for GCA were especially strong within the
23 Major Histocompatibility Complex (MHC) (See Fig-S4). Notably, anorexia nervosa and ADHD
24 show the opposite pattern of local genetic correlations within the locus relative to other
25 psychopathological traits.

1 Local genetic correlation profiles were further parsed to better understand pleiotropic
2 patterns across cognitive function and psychopathology. UMAP dimension reduction, followed
3 by clustering (using density-based scanning, DBscan) was applied to the matrix of local genetic
4 correlations (17 psychopathologic traits x 2330 LD-independent regions, excluding the MHC) to
5 identify clusters of LD-independent regions with distinct pleiotropic patterns for GCA or NCS
6 across the various psychopathologic traits (see Online Methodology). Resulting clusters of LD-
7 independent regions were termed “Meta-Loci”. We identified 15 meta-loci for GCA and 12 meta-
8 loci for NCS (Fig-2a/b; see Supplementary Information for further details of the DBscan
9 method). Critically, the extracted meta-loci were not simply a function of highly localized effects
10 but were distributed across the genome (Fig-2c/d). Each meta-locus encompassed between 7 -
11 1125 (GCA; median = 28) and 8 - 1926 (NCS, median=16) LD-independent segments of the
12 genome (total length 6.96 - 1340.79 Mb, median=34.24 Mb for GCA; 8.70 - 2290.00 Mb,
13 median=23.19 Mb for NCS). For both GCA and NCS, the first meta-locus extracted (GCA-1 and
14 NCS-1, respectively, indicated in grey in Fig-2), consists of all genomic segments that could not
15 be further clustered by the DBscan search without being listed as outliers, such that the entire
16 genome is accounted for. We prioritized meta-loci based on their relative contribution to their
17 total heritability: 8 meta-loci for GCA explained 95% of the total heritability of GCA (GCA-
18 1,3,4,5,6,7,8,12); similarly, 4 meta-loci for NCS were prioritized (NCS-1,2,3,5) (Fig-S3, Table-
19 S4b/S5b).

20 The pattern of local genetic correlation for each of the prioritized meta-loci is displayed in
21 Fig-3 (distribution for all meta-loci displayed in Fig-S5 & point estimates of mean local genetic
22 correlation Z-scores displayed in Fig-S6). Schizophrenia and bipolar disorder showed similar
23 local genetic correlation profiles for the prioritized NCS meta-loci but were differentiated on the
24 GCA meta-loci. As expected, Schizophrenia showed strong trends of negative genetic
25 correlation in most GCA meta-loci, with positive local genetic correlations observed in only 2 of
26 the 8 prioritized GCA meta-loci. Notably, attention-deficit/hyperactivity disorder (ADHD) was

1 negatively associated across all GCA and NCS meta-loci, whereas obsessive compulsive
2 disorder (OCD) was positively associated with both cognitive dimensions across all meta-loci.
3 Affective and anxiety related traits showed expected broad negative associations with both
4 cognitive dimensions, but appeared differentiated on specific meta-loci e.g., GCA-3. Autism
5 spectrum disorder (ASD) demonstrated a relatively unique pattern of relationships, with
6 relatively modest effects across most meta-loci, except for a positive genetic correlation with
7 GCA at meta-locus GCA-7, and a negative correlation at NCS-3.

8 It was important to determine if the pattern of results was driven by socioeconomic
9 status, given that prior literature³³ has indicated that educational attainment exhibits shared
10 biology with socioeconomic status. We carried out local genetic correlations of GCA and NCS
11 and the Townsend Social Deprivation Index via ρ -HESS, in a similar manner detailed above. As
12 displayed in Fig-S7/Table-S7, most meta-loci showed negligible association.

13 **Stage 3: Functional annotation and gene prioritization for characterizing biology** 14 **underlying GCA and NCS Meta-Loci**

15 In order to characterize each meta-locus, we first applied a series of gene-based and
16 TWAS methods to the GCA and NCS GWAS summary statistics (gene-based results: Table-
17 S8; TWAS results: Table-S9/S10/S11). As an initial loose filter, genes were ranked based on
18 multiple complementary transcriptome analyses (Table-S12), applying a 50th percentile cut off
19 for gene association p-values; any genes with no evidence of transcriptomic association to the
20 cognitive phenotypes were removed from further consideration. PoPs gene prioritization
21 scores³² (Table-S13) for each remaining gene within a given meta-locus were then subjected to
22 gene set analyses, using annotations obtained from the GO Gene Sets (Cellular Component,
23 Molecular Function and Biological Process) included in the Molecular Signature Database v
24 7.2³⁴. Detailed descriptions of the parameters applied to each of the methods are described in
25 the Methodology section. Gene set analyses were carried out via three methods [Gene Set
26 Enrichment Analysis (GSEA²⁹), WebGestalt³⁰, and GENE2FUNC (part of the Functional

1 Mapping and Annotation of Genetic Association – FUMA – pipeline³¹]; we conservatively
2 required agreement (Bonferroni-corrected significance) from all three methods in order to
3 designate a gene set as significant for a given meta-locus.

4 The gene set analysis described above identified a total of 62 gene sets underlying 6
5 GCA meta-loci (GCA-1,3,4,5,6,8) and 35 gene sets underlying 3 NCS meta-loci (NCS-1,2,5)
6 (Fig-4a, and Table-S14); the remaining meta-loci had no gene sets meeting our strict
7 significance criteria. In Fig-4, several large neurobiological gene sets which were implicated
8 across both GCA and NCS meta-loci are colored red (Fig-4a/b) and those that overlapped
9 multiple GCA meta-loci (but no NCS meta-loci) are colored orange (Fig-4c). The remaining gene
10 sets, which are specific to individual GCA or NCS meta-loci, are colored gray.

11 Focusing on gene sets specific to individual meta-loci might allow more targeted parsing
12 of the biological mechanisms underlying the relationship between psychopathology and
13 cognition. There were 34 such gene sets across GCA meta-loci, sharing a predominately
14 neurodevelopmental theme as elaborated in Table 1 (e.g., GCA-1: Neurodevelopment/
15 Neuroplasticity; GCA-4: Neural Connectivity/Differentiation). By contrast, the 24 gene sets
16 specific to individual NCS meta-loci shared a predominant theme relating to synaptic
17 transmission (e.g., NCS-1: Synaptic Activity/Neurotransmission; NCS-2: Synaptic Transmission
18 /Potentiation; NCS-5: GABA-ergic Synaptic Transmission) (Table 2). Pathway definitions for
19 each prioritized meta-locus, as a function of psychopathologic correlations are summarized in
20 Fig-S8.

21 There were 906 GCA and 493 NCS “driver” genes identified from the GSEA analysis.
22 The reduced list of genes was further examined if they were potentially actionable in terms of
23 having encoding proteins of putative drug targets, using chemoinformatic annotations provided
24 by Finan et al.³⁵ (Table-S15). Tier 1 druggable genes (i.e., genes with current evidence of
25 having existing compounds that are FDA approved and being utilized for various indications),
26 summarized alongside other meta-locus information, are listed in Table 1 and Table 2. Several

1 gene families with known druggable components were identified, including several
2 glutamatergic and ephrin receptor genes identified within GCA meta-loci. Most notably, the
3 NCS-5 meta-locus implicates multiple GABA receptor genes across disparate chromosomal
4 regions, suggesting the possibility that GABA-ergic treatment approaches may enhance NCS
5 while simultaneously ameliorating the correlated affective and anxiety symptoms that load on
6 this meta-locus.

7 The dichotomy between neurodevelopment and synaptic pathways underlying GCA and
8 NCS, respectively, suggested an additional hypothesis to be tested. If GCA was primarily driven
9 by neurodevelopmental genes, we would expect to see gene expression profiles that are active
10 prenatally or early in the lifespan, whereas for NCS, genes responsible for synaptic transmission
11 are likely to be expressed later in postnatal life. Leveraging the BrainSpan dataset³⁵ which
12 characterized brain expression profiles at various developmental stages, we tested if GCA and
13 NCS driver genes might demonstrate differential expression across the lifespan. By fitting a
14 linear mixed model, with individual as random effect, cognitive phenotypes and time as fixed
15 effects and sex as a covariate, we found developmental differences between GCA and NCS
16 driver genes (for further details, see Online Methodology and Supplementary Information).
17 Setting the null model simply with time as predictor, we observed a significant improvement in
18 model fit when an indicator of the cognitive phenotype (GCA vs. NCS) for the driver gene was
19 included in the model ($\chi^2 = 1059.6$, $df = 2$, $p = 8.14e-231$). A main effect for differences in GCA
20 and NCS genes was found ($t = 10.64$, $df = 7$, $p = 1.42e-5$), and an interaction between GCA and
21 NCS genes with time ($t = 28.28$, $df = 7$, $p = 1.78e-8$) was also observed (Fig-5, Fig-S9). As
22 hypothesized, genes harbored within NCS-meta loci were expressed predominantly in early
23 adulthood and adulthood, whereas GCA-meta loci genes were expressed pre-natally.

24

25

1 Discussion

2 Numerous recent studies have demonstrated shared genetic underpinnings amongst
3 multiple forms of psychopathology, as well as between psychiatric and cognitive
4 phenotypes^{10,11,23,37}. The present study was designed to parse this genetic overlap into
5 separable biological pathways with specific psychiatric and cognitive subcomponents. The
6 present study extends our previous work, in which we leveraged cognitive pleiotropy to
7 differentiate early neurodevelopmental mechanisms from synaptic function in the
8 etiopathogenesis of schizophrenia¹⁵. Here, we demonstrated that prenatal neurodevelopmental
9 mechanisms shape the relationship between general cognitive ability and multiple forms of
10 psychopathology, while synaptic pathways expressed later in life underlie the paradoxical
11 association between enhanced non-cognitive skills and greater risk for psychotic disorders.
12 Prior studies of pleiotropy^{10,11,14} have implicated neurodevelopmental and synaptic pathways as
13 shared components underpinning multiple psychiatric disorders, but without the ability to
14 differentiate these mechanisms (Table-S16). By contrast, gene set analysis of individual trait
15 GWAS in psychiatry (Table-S16) have either failed to identify one^{38,39} or both⁴⁰ of these
16 pathways, or have implicated these pathways without directionality or subsetting^{41,42}

17 Within the broad neurodevelopmental and synaptic themes, we were able to further specify
18 individual gene sets with distinct patterns of association to psychiatric and cognitive phenotypes.
19 These discoveries were made possible by our novel construction of “meta-loci”, defined by the
20 concatenation of LD-independent genomic regions with shared patterns of local genetic
21 correlations across phenotypes. Within meta-loci, we further identified genes that may serve as
22 actionable targets, or at least actionable entry points into relevant biological pathways, for
23 psychiatric and nootropic drug repurposing (Table 1 and Table 2); several notable examples are
24 described below.

1 The GCA-8 meta-locus spans 55 LD-independent regions, covering the genomic length
2 of 73Mb, and is characterized by “Axon Development” and related gene sets. We found that this
3 meta-locus shows the strongest pleiotropy with schizophrenia, but also implicates bipolar
4 disorder and a series of affective phenotypes. Notable driver genes in this meta-locus included
5 *DRD2*, *NCAM1*, *DCC*, *CACNA1D*, and *SLC4A10*, among many known genes that implicate
6 psychopathology. In their most recent report, the PGC Cross Disorder group reported that the
7 *DCC* region was most pleiotropic in psychopathology. The *DCC* gene was a known master
8 regulator of early neurodevelopmental biology via interactions with netrin-1 and draxin, and has
9 been implicated in developmental processes of white matter tracts in the brain^{43,44}. In the report,
10 the authors suggested that the *DCC* is likely to affect childhood developmental disorders such
11 as ADHD and ASD, which they have demonstrated to cluster together. Importantly, our
12 evidence showed that similar processes are not limited to neurodevelopmental disorders but are
13 implicated in other psychopathological phenotypes and are linked through a shared effect on
14 general cognitive ability.

15 By contrast, the NCS-5 locus showed the strongest trend for adulthood spatial-temporal
16 gene expression and the driver genes were primarily inhibitory GABA genes, several of which
17 were previously implicated in major depressive disorder⁴⁵ and post-traumatic stress disorder⁴⁶.
18 Similarly, driver genes at the NCS-2 meta-locus, characterized by “Synaptic Transmission” and
19 related gene sets, showed predominant postnatal/adulthood expression. This meta-locus shows
20 strong pleiotropy with psychotic disorders (schizophrenia and bipolar disorder); to lesser extent
21 positive local genetic correlations were also observed in affective disorders and anxiety. *AKAP6*,
22 *CADM3*, and *PDE4B* were notable genes implicated in both schizophrenia and bipolar in earlier
23 reports⁴⁷⁻⁴⁹. At the same time, novel associations were observed to *PPARG* and *HMGCR*, both
24 of which are critical to metabolism and energy utilization, possibly suggesting an unexplored
25 mechanism of neuronal function. Given the counterintuitive pleiotropy at this locus, with greater

1 non-cognitive skills associated with increased risk for psychopathology, such mechanistic leads
2 may prove fruitful for further experimentation.

3 The meta-locus approach also extends and refines prior work on the structure of
4 psychopathology. Recently, both the PGC Cross Disorder Group¹⁰ as well as Grotzinger and
5 colleagues³⁶ utilized global genetic correlations to show converging evidence for several latent
6 factors underlying current nosological constructs in psychiatry: (i) Psychosis factor
7 (schizophrenia and bipolar disorder); (ii) Neurodevelopmental factor (alcohol use, ADHD, ASD,
8 PTSD); (iii) Compulsive factor (anorexia nervosa, OCD, and Tourette's Syndrome); and (iv)
9 Internalizing factor (MDD, anxiety disorders). In this report, schizophrenia and bipolar disorder
10 demonstrated similar patterns of pleiotropy globally, but distinctions were observed at several
11 meta-loci (e.g., GCA-5 and GCA-7), where schizophrenia more closely resembled major
12 depression and other affective phenotypes. Overall, schizophrenia harbors LD-independent
13 segments strongly associated with either GCA or NCS beyond that of other psychopathological
14 phenotypes. Though the genetic correlation between schizophrenia and cognitive dimensions
15 appears modest ($r_g \approx -0.2$) at the global level, this may reflect mutually negating effects at the
16 regional level across the genome. Broad pleiotropic profiles for GCA/NCS within ADHD, PTSD,
17 MDD, Anxiety disorder and Tourette's Syndrome support cross factor loadings observed by
18 Grotzinger and colleagues³⁶ within "Internalizing" and "Neurodevelopmental factors". However,
19 in the present study, ADHD and ASD exhibited quite different local genetic correlation profiles
20 across most meta-loci – ADHD being negatively correlated across all meta-loci for GCA and
21 NCS, while ASD demonstrated null or positive correlations with both GCA and NCS. We also
22 demonstrated that the strong positive genetic correlations reported previously between the
23 "Compulsive factor" and Education Attainment was likely accounted for by higher-than-expected
24 genetic correlations between OCD and Non-Cognitive Skills.

25 Although the MHC region was excluded from most of the downstream work due to its
26 complicated LD patterns, it is worth mentioning that it showed stronger GCA-psychopathology

1 local genetic correlations than other genome regions. Also noteworthy is that Anorexia Nervosa
2 and Tourette's syndrome showed opposite local genetic correlations within the MHC region
3 relative to all other GCA/NCS - psychopathology trait pairs. Evidence points to the MHC region
4 as potentially a vital aspect of etiopathogenesis in psychopathology. The MHC region harbors
5 the strongest genome-wide signal to date for psychotic and affective disorders. It also harbors a
6 known synaptic pruning mechanism as part of the *C4* complex. The strong association with
7 cognitive ability makes MHC a candidate region for further extensive investigation.

8 *Conclusions*

9 To conclude, we have leveraged the phenomenon of pleiotropy between cognitive
10 dimensions and psychopathology to dissect underlying biological mechanisms. We
11 demonstrated that beyond broad genome-wide relationships, specific regions of the genome are
12 likely to harbor biological processes that act differentially across psychopathological traits. The
13 phenomenon is unveiled through the combined local genetic correlation analysis and "meta-
14 locus" strategy adopted within this report. Follow up efforts such as increasing the power of
15 univariate traits and the accuracy of reference panels may increase the resolution of such
16 approaches. Also, it seems necessary to devise a way in which latent factors such as Non-
17 Cognitive Skills derived purely from genetic correlations and GWAS summary statistics could be
18 operationalized and measured tangibly in clinical populations. Finally, it should be noted that the
19 present study utilizes well-powered, currently available GWAS that are based on common
20 genetic variation and community sampling; in the future, well-powered family based GWAS may
21 ultimately provide more accurate estimates of genetic effects, including rare variation⁵⁰,
22 independent of "genetic nurture"⁵¹. Nonetheless, the results reported currently underscores the
23 need to rethink how rapidly available multivariate data within cognitive and psychiatric genomics
24 could be leveraged to uncover more in-depth biological insights beyond simply pointing towards
25 neuronal mechanisms.

26

1 **Online Methodology**

2 *Genome-Wide Summary Statistics Quality Control and Harmonization*

3 The first stage of the analysis involved the curation of the GWAS summary statistics of
4 cognitive features and psychopathological traits. We examined the recent literature and selected
5 GWAS of 19 cognition related traits as well as 17 psychopathological traits previously known to
6 be with associated with cognitive function (See Figure 1, Supplementary Table 1). All summary
7 statistics were harmonized based on alleles and frequencies in the 1000 Genomes phase 3 v5
8 reference⁵². Variants with allele frequencies less than 0.5% and imputation quality less than 0.3
9 were excluded from the analysis. We also performed allele alignment to resolve strand and
10 allele flips. All summary statistics were aligned based on the A1 allele (reference allele) of the
11 1000G reference panel⁵². Variants that could not be aligned to the 1000 genomes were
12 excluded. In addition, we excluded ambiguous variants with an allele frequency difference of
13 greater than 0.35 from the reference panel and non-ambiguous variants with an allele frequency
14 difference of 0.15 difference from the reference. Further details are reported in Supplementary
15 Table 1.

16 *Global Genetic Correlations and Clustering Analysis*

17 We carried out global genetic correlation analysis between the set of 19 cognitive traits
18 and 17 psychiatric traits via LD score regression^{21,22} implemented in Genomic SEM¹¹. We
19 followed standard procedures within Genomic SEM as a wrapper for extracting the global
20 genetic covariance matrix (<https://github.com/GenomicSEM/GenomicSEM/wiki/>). GenomicSEM
21 performs initial data alignment to the HAPMAP3 SNPs and conducts pairwise LD score
22 regression with each pair of input phenotypes for the estimation of global genetic correlations.
23 The global genetic correlations between all cognitive and psychopathological traits were
24 organized into a 19 x 17 matrix and subsequently used as the input for Principal Components
25 Analysis (PCA) and various clustering analyses reported subsequently. The first two principal

1 components were extracted from the 19 x 17 matrix representing the relationship between each
2 psychopathology trait and the top two cognitive dimensions (Table-S2). The loading of each
3 cognitive trait on the principal component was estimated by performing bivariate Pearson
4 correlation between the genetic correlation profile of each cognitive trait and each principal
5 component (columns U and V in Table-S2). The dissimilarity matrix estimated based on
6 Euclidean distance from the global genetic correlation matrix was used to generate the
7 partitioned k-medoid clusters. The clustering procedures were carried out using `fviz_cluster()` in
8 R as part of the “factoextra” package
9 (<https://www.rdocumentation.org/packages/factoextra/versions/1.0.3>)

10 After closely reviewing the global genetic correlation profiles for cognitive features and
11 psychopathological traits described above, GCA and NCS emerged as two candidate traits that
12 were fairly well differentiated across clusters of psychopathological traits. Partitioning cluster
13 analysis (K-medoids and K-means) and hierarchical clustering were employed to examine the
14 latent genome-wide genetic architecture of psychopathological traits in relation to cognitive
15 features. The strategy for the clustering analysis was as follows - we started with a 2-cluster
16 solution, and gradually increased the number of clusters until any of the method reached a
17 singleton cluster. Across methodologies, a 7-cluster solution resulted in a singleton cluster,
18 hence we set the maximum number of clusters to 6 and the minimum number of clusters 2. To
19 select the optimal cluster solution that best represents the data, the procedures detailed in
20 Akhanli and Hennig⁵³ were used, which compared model fitting statistics across clustering
21 solutions, using R packages `cqcluster.stats()` and `clusterboot()` with default parameters. The best
22 fitting cluster solution was then selected for visualization.

23

24

25

26

1 *Data Reduction of Cognitive Features*

2 We applied the GWAS-by-subtraction¹⁶ procedures developed recently to the latest and
3 largest GCA GWAS²⁰ to derive a updated NCS GWAS. We used the Education Attainment
4 GWAS without 23andMe data (N = 495,552), which was slightly smaller compared to the
5 Education Attainment GWAS used in Demange et al.¹⁶ (N = 565,884). However, our GCA
6 GWAS (N = 373,617) was much more powerful than the one reported previously¹⁶ (N =
7 257,841). Nonetheless, LD score regression checks indicated that both GCA and NCS GWAS
8 were highly genetically concordant to previous results ($R_g = 1$).

9

10 *Identify pleiotropy between Cognition and Non-Cognitive Skills and Psychopathological Traits* 11 *via Local Genetic Correlations*

12 Local genetic correlation analysis was conducted via ρ -HESS¹⁸. On the advice of the
13 HESS developers, we used a conservative approach that specified full sample overlap
14 between all trait pairs to handle potential unknown sample overlaps. Independent LD segments
15 were estimated using the 1000 genome phase 3 reference panel via LDetect^{54,55}. There were a
16 total of 2353 independent LD segments where local genetic correlations for GCA and NCS were
17 carried out vis-a-vis psychopathological traits.

18

19 *Identification of Meta-loci via Density-Based Clustering*

20 In order to uncover homogeneous local genetic correlation patterns (between GCA and
21 psychopathological traits and between NCS and psychopathological traits) across the genome,
22 we carried out density-based clustering⁵⁶ in two steps. First, we performed a data reduction step
23 via UMAP⁵⁷ for all pairs of the 19 cognitive features and 17 psychopathological traits across
24 2353 independent LD segments. Thereafter, density-based scanning was carried out on the first
25 two UMAP components to identify sets of LD independent regions that had homogeneous local
26 genetic correlation profiles between GCA/NCS and psychopathology; these sets of LD

1 independent regions were termed as meta-loci. The density-based scan was carried out by first
2 setting the minimum number of LD-independent regions per set as 5 and setting the
3 neighborhood radius to a sufficiently large value such that only one cluster existed. We then
4 reduced the neighborhood radius in increments of 0.1, until the density-based scan yielded a
5 cluster solution d with a singleton LD region outlier that could not be clustered. The cluster
6 solution $d - 1$ was used (see Supplementary Information for further details).

7

8 *Functional Prioritization of Putative Genes within Meta-loci*

9 To obtain biological insights from the meta-loci we identified, we first prioritized genes
10 within the broad genomic regions based on their possible associations with cognitive
11 phenotypes. We applied a series of gene-based statistical approaches to the GCA and NCS
12 summary statistics derived from the Genomic SEM procedure; these included MAGMA²⁸,
13 PoPs³², S-Predixcan⁵⁸, SMR/HEIDI^{59,60}, and FOCUS⁶¹. As our goal was to identify relevant
14 genes that may include those below the level of strict genome-wide significance, our initial filter
15 was loose: any gene that was below the 50th percentile of statistical significance for *all* of the
16 above methods was eliminated from further consideration.

17 For the MAGMA gene-based analysis (v1.08²⁸), gene boundaries definitions and the
18 1000 genomes phase 3 reference panel were provided with the MAGMA package
19 (<https://ctg.cncr.nl/software/magma>). PoPs allows additional weighting of MAGMA gene-based
20 results, by leveraging the extraction of gene features from comparable genes across all other
21 chromosomes; we applied default parameters as described in Weeks et al.³². S-PrediXcan
22 estimates gene-based effects by re-weighting SNP effects with eQTL effects derived from GTEx
23 v8 brain tissue expression data (with LD information derived from the 1000 Genomes phase 3
24 reference panel). SMR/HEIDI leverages on Mendelian Randomization approaches for gene
25 prioritization; each SNP was treated as an instrument variable, with GCA or NCS as an
26 outcome. Brain expression eQTL data from Brain-eMeta²⁵, and PsychENCODE^{26,27} reference

1 datasets are the intermediate variables. As part of the SMR analysis, HEIDI attempts to
2 deconvolve the association to establish if the relationship between SNP, gene expression, and
3 cognitive phenotype are indeed due to modelled causality, or based simply on linkage; SNPs
4 with significant evidence of linkage effects were eliminated from further consideration. FOCUS
5 (Fine-Mapping of Causal Variants)⁶¹ uses GCA/NCS GWAS summary statistics, expression
6 prediction weights (as estimated from GTEx v8 brain expression reference panels), and LD
7 among all SNPs to estimate the probability for any given set of genes to explain transcriptome
8 association. Two sets of gene ranks emanated from the procedures carried out (i)
9 MAGMA/PoPs (ii) S-PrediXcan/SMR/HEIDI. To prioritize genes, the 50th percentile rank cutoff
10 for (i) & (ii) was stipulated. The final set of filtered genes for gene set analysis was the union of
11 (i) & (ii) and the union of all credible genes identified by FOCUS.

12

13 *Elucidating Biological Mechanisms underlying Meta-Loci via Gene Set Analysis*

14 Three separate gene set analysis approaches were then applied to the remaining,
15 filtered gene list for each meta-locus (i) Broad Institute Gene Set Enrichment Analysis
16 (GSEA^{29,62}) (ii) WebGestalt³⁰ (iii) FUMA GENE2FUNC³¹. GO ontologies within the Molecular
17 Signature Database 7.2 (Biological Processes, Molecular Function and Cellular Component^{34,63})
18 were used as gene-set analysis annotations. Default parameters were used for GSEA and
19 WebGestalt. Both GSEA and WebGestalt require gene scores to be included in the gene set
20 analysis; we utilized the gene prioritization scores derived from PoPs, with an inverse rank
21 transformation applied to the POPs data such that the lowest value of the gene prioritization
22 scores was 0.

23 Of the three gene set analysis methods, GSEA was arguably the most robust. GSEA
24 first walks down the ranked list of genes, increasing a running-sum statistic when a gene is in
25 the gene set and decreasing it when it is not. The enrichment score is the maximum deviation
26 from zero encountered during that walk. Details of the GSEA methodology has been reported

1 elsewhere²⁹. This is followed by WebGestalt⁶⁴ which runs a modified version of GSEA via a web
2 portal. Both GSEA and WebGestalt allow permutation testing for the selection of gene sets
3 which in this case was set to $n = 1000$ permutations. GSEA analyses were based on the
4 “preranked” procedure, where enrichment scores were normalized. Default filter parameters for
5 minimum (> 15) and maximum (< 500) gene set size were used. MSigDb version 7.2 gene set
6 definitions³⁴ for Gene Ontology were used as indicated above. WebGestalt uses a more
7 specifically curated “noRedundant” set of Gene Ontologies based on the 2017 data freeze of the
8 MSigDb. In addition, the method relaxes gene set sizes to permit minimum (> 5) and maximum
9 (< 2000). We set significance to the top 50 gene sets to be extracted for WebGestalt. FUMA
10 GENE2FUNC uses a hypergeometric approach to gene selection which relies only on the
11 overrepresentation of gene symbols for the identification of gene sets. For FUMA GENE2FUNC
12 we set a minimum of 3 genes per gene set and $FDR < 0.05$ for a gene set to be significant. To
13 select candidate gene sets for each meta-locus, we required a strict “consensus” approach. We
14 deem a gene set as “candidate” only if all three methods identify the same gene set.

15 In addition to an enrichment score for the gene set, GSEA also generates a set of genes
16 that deviates most from the null during the estimation of the gene set enrichment score. These
17 genes are defined as “driver” genes, because they contribute most to the enrichment of a given
18 gene set. For the purpose of the current report, the driver genes were thus most interesting, as
19 they are likely to be the core genes within a biological pathway.

20

21 *Follow-up Annotations and Analysis of Driver Genes identified within GCA and NCS Meta-Loci*

22 To further understand how driver genes from each gene set might be expressed over
23 time, we performed analyses via the BrainSpan database (based on RNA-Seq Gencode v10
24 normalized expression values summarized to genes in brain tissues from donors of different
25 ages, ranging from pre-natal to adulthood)³⁵. For our analysis, the time of tissue sampling in the
26 BrainSpan database was converted to weeks. All post-natal stages were converted to weeks

1 using $Weeks = Years * 52 Weeks + 37 Gestational Weeks$. We used linear mixed effects models
2 to investigate if there were significant differences in the gene expression profiles over time
3 between GCA vs. NCS driver genes. The null model included Individual subjects as random
4 effects, and *Time (Weeks)* as a fixed effect predictor; sex was modeled as a covariate. The
5 alternative model additionally included the *Time (Weeks) x Trait (GCA vs. NCS)* interaction
6 terms as fixed effect predictors.

7 Lastly, we examined the potential druggability of the driver genes identified by earlier
8 functional prioritization and gene sets. We extracted data for gene druggability (2047 genes)
9 from Finan et al.³⁵ which included Tier 1, 2, 3A and 3B druggability as well as small molecule
10 druggability, druggability by enzymes or mono-clonal antibodies and drug absorption
11 characteristics. The druggability information was merged with the information available for driver
12 genes identified in the current study. We then further filtered the final list of genes by Tier 1
13 druggability for display and reporting.

14

1 **Acknowledgments:** Research reported in this publication was supported by the National
2 Institute of Mental Health of the National Institutes of Health under award number R01
3 MH117646 (TL, Principal Investigator). The content is solely the responsibility of the
4 authors and does not necessarily represent the official views of the National Institutes of
5 Health. WDH and CX are supported by a Career Development Award from the Medical
6 Research Council (MRC) [MR/T03052/1] for the project titled: "From genetic sequence to
7 phenotypic consequence: genetic and environmental links between cognitive ability,
8 socioeconomic position and health."
9

10 **Author Contributions:** ML ran the analyses and drafted the manuscript. ML and TL conceived
11 the idea and designed the study. TL supervised the drafting and analyses reported in the
12 manuscript. TG, C-YC, WDH, CX, and RT provided guidance on data analytic approaches.
13 HH, AKM, and HR provided guidance on interpretation of the results. C-YC, ASH, DFL, JG,
14 and MBS provided summary statistics for various GWAS entered into the study. All authors
15 read and provided scientific feedback and participated in finalizing the draft of the
16 manuscript.
17

18 **Competing Interests:** Three authors (C-YC, RT, and HR) are employees of Biogen. All other
19 authors have no competing interests pertaining to this study.
20

21 **Data Availability:** GWAS summary statistics for GCA and NCS will be made publicly available
22 upon acceptance of the manuscript. Publicly available GWAS summary statistics for
23 psychopathology traits can be obtained from the respective research groups/consortia
24

25 **Code Availability:** Wrapper for performing local genetic correlations reported in the current
26 report will be made publicly available upon acceptance of the manuscript.

- 1 **Statement of ethics:** GWAS summary statistics utilized in the current study are publicly
- 2 available except for the UK Biobank GWAS summary statistics for cognitive phenotypes and the
- 3 Townsend Index for Social Deprivation which was approved for use under UK Biobank
- 4 Approved Application ID 26041 (“Large-Scale Sequencing in the UK Biobank to Facilitate Gene
- 5 Discovery, Genome Sciences, and Precision Medicine”).

1 **References**

- 2 1. Van Rheenen, T. E., Lewandowski, K. E., Lipschitz, J. M. & Burdick, K. E. Conducting
3 clinical studies targeting cognition in psychiatry: guiding principles and design. *CNS Spectr.*
4 **24**, 16–21 (2019).
- 5 2. Green, M. F. Cognitive impairment and functional outcome in schizophrenia and bipolar
6 disorder. *J. Clin. Psychiatry* **67 Suppl 9**, 3–8; discussion 36-42 (2006).
- 7 3. Bora, E. *et al.* Cognitive deficits in youth with familial and clinical high risk to psychosis: a
8 systematic review and meta-analysis. *Acta Psychiatr. Scand.* **130**, 1–15 (2014).
- 9 4. Lam, M. *et al.* Longitudinal Cognitive Changes in Young Individuals at Ultrahigh Risk for
10 Psychosis. *JAMA Psychiatry* **75**, 929–939 (2018).
- 11 5. Karalunas, S. L. *et al.* Overlapping and Distinct Cognitive Impairments in Attention-
12 Deficit/Hyperactivity and Autism Spectrum Disorder without Intellectual Disability. *J.*
13 *Abnorm. Child Psychol.* **46**, 1705–1716 (2018).
- 14 6. Reichenberg, A. *et al.* Static and dynamic cognitive deficits in childhood preceding adult
15 schizophrenia: a 30-year study. *Am. J. Psychiatry* **167**, 160–169 (2010).
- 16 7. Paus, T., Keshavan, M. & Giedd, J. N. Why do many psychiatric disorders emerge during
17 adolescence? *Nat. Rev. Neurosci.* **9**, 947–957 (2008).
- 18 8. Gottesman, I. I. & Gould, T. D. The endophenotype concept in psychiatry: Etymology and
19 strategic intentions. *Am. J. Psychiatry* **160**, 636–645 (2003).
- 20 9. Lencz, T. *et al.* Molecular genetic evidence for overlap between general cognitive ability
21 and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT).
22 *Mol. Psychiatry* **19**, 168–174 (2014).
- 23 10. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships,
24 Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-
25 1482.e11 (2019).

- 1 11. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the
2 multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525 (2019).
- 3 12. Chesmore, K., Bartlett, J. & Williams, S. M. The ubiquity of pleiotropy in human disease.
4 *Hum. Genet.* **137**, 39–44 (2018).
- 5 13. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic
6 correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–
7 581 (2019).
- 8 14. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex
9 traits. *Nat. Genet.* **51**, 1339–1348 (2019).
- 10 15. Lam, M. *et al.* Pleiotropic Meta-Analysis of Cognition, Education, and Schizophrenia
11 Differentiates Roles of Early Neurodevelopmental and Adult Synaptic Pathways. *Am. J.*
12 *Hum. Genet.* **105**, 334–350 (2019).
- 13 16. Demange, P. A. *et al.* Investigating the genetic architecture of noncognitive skills using
14 GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).
- 15 17. The Brainstorm Consortium *et al.* Analysis of shared heritability in common disorders of the
16 brain. *Science* **360**, eaap8757 (2018).
- 17 18. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives
18 Insights into the Shared Genetic Architecture of Complex Traits. *Am. J. Hum. Genet.* **101**,
19 737–751 (2017).
- 20 19. Hill, W. D. *et al.* A combined analysis of genetically correlated traits identifies 187 loci and a
21 role for neurogenesis and myelination in intelligence. *Mol. Psychiatry* **24**, 169–181 (2019).
- 22 20. Lam, M. *et al.* Identifying nootropic drug targets via large-scale cognitive GWAS and
23 transcriptomics. *Neuropsychopharm* 1–14 (2021).
- 24 21. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits.
25 *Nat. Genet.* **47**, 1236–1241 (2015).

- 1 22. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity
2 in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
- 3 23. Hatoum, A. S. *et al.* Genome-Wide Association Study of Over 427,000 Individuals
4 Establishes Executive Functioning as a Neurocognitive Basis of Psychiatric Disorders
5 Influenced by GABAergic Processes. *Cold Spring Harbor Laboratory* 674515 (2020)
6 doi:10.1101/674515.
- 7 24. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* vol.
8 550 204–213 (2017).
- 9 25. Qi, T. *et al.* Identifying gene targets for brain-related traits using transcriptomic and
10 methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018).
- 11 26. Gandal, M. J. *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia,
12 and bipolar disorder. *Science* **362**, (2018).
- 13 27. Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the
14 human brain. *Science* **362**, eaat6484 (2018).
- 15 28. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set
16 Analysis of GWAS Data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
- 17 29. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for
18 interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–
19 15550 (2005).
- 20 30. Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z. & Zhang, B. WebGestalt 2019: gene set analysis
21 toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **47**, W199–W205 (2019).
- 22 31. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and
23 annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- 24 32. Weeks, E. M. *et al.* Leveraging polygenic enrichments of gene features to predict genes
25 underlying complex traits and diseases. *medRxiv* (2020).

- 1 33. Hill, W. D. *et al.* Genome-wide analysis identifies molecular systems and 149 genetic loci
2 associated with income. *Nat. Commun.* **10**, 5741 (2019).
- 3 34. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set
4 collection. *Cell Syst.* **1**, 417–425 (2015).
- 5 35. Finan, C. *et al.* The druggable genome and support for target identification and validation in
6 drug development. *Sci. Transl. Med.* **9**, eaag1166 (2017).
- 7 36. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–
8 206 (2014).
- 9 37. Grotzinger, A. D. *et al.* Genetic architecture of 11 major psychiatric disorders at
10 biobehavioral, functional genomic, and molecular genetic levels of analysis. *bioRxiv* (2020)
11 doi:10.1101/2020.09.22.20196089.
- 12 38. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention
13 deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
- 14 39. Mullins, N. *et al.* Genome-wide association study of over 40,000 bipolar disorder cases
15 provides new insights into the underlying biology. *medRxiv* (2020)
16 doi:10.1101/2020.09.17.20187054.
- 17 40. Nagel, M. *et al.* Meta-analysis of genome-wide association studies for neuroticism in
18 449,484 individuals identifies novel genetic loci and pathways. *Nat. Genet.* **50**, 920–927
19 (2018).
- 20 41. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke, S., Walters,
21 J. T. R. & O'Donovan, M. C. Mapping genomic loci prioritises genes and implicates synaptic
22 biology in schizophrenia. *Genetic and Genomic Medicine* (2020)
23 doi:10.1101/2020.09.12.20192922.
- 24 42. Levey, D. F. *et al.* GWAS of depression phenotypes in the Million Veteran Program and
25 meta-analysis in more than 1.2 million participants yields 178 independent risk loci.
26 *medRxiv* (2020) doi:10.1101/2020.05.18.20100685.

- 1 43. Bendriem, R. M. & Ross, M. E. Wiring the Human Brain: A User's Handbook. *Neuron* **95**,
- 2 482–485 (2017).
- 3 44. Liu, Y. *et al.* Structural Basis for Draxin-Modulated Axon Guidance and Fasciculation by
- 4 Netrin-1 through DCC. *Neuron* **97**, 1261-1267.e4 (2018).
- 5 45. Hall, L. S. *et al.* Genome-wide meta-analyses of stratified depression in Generation
- 6 Scotland and UK Biobank. *Transl. Psychiatry* **8**, 9 (2018).
- 7 46. Ashley-Koch, A. E. *et al.* Genome-wide association study of posttraumatic stress disorder in
- 8 a cohort of Iraq-Afghanistan era veterans. *J. Affect. Disord.* **184**, 225–234 (2015).
- 9 47. Goes, F. S. *et al.* Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am.*
- 10 *J. Med. Genet. B Neuropsychiatr. Genet.* **168**, 649–659 (2015).
- 11 48. Ikeda, M. *et al.* Genome-wide association study detected novel susceptibility genes for
- 12 schizophrenia and shared trans-populations/diseases genetic effect. *Schizophr. Bull.* **45**,
- 13 824–834 (2019).
- 14 49. Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar
- 15 disorder. *Nat. Genet.* **51**, 793–803 (2019).
- 16 50. Hill, W. D. *et al.* Genomic analysis of family data reveals additional genetic effects on
- 17 intelligence and personality. *Mol. Psychiatry* **23**, 2347–2362 (2018).
- 18 51. Howe, L. J. *et al.* Within-sibship GWAS improve estimates of direct genetic effects. *bioRxiv*
- 19 (2021) doi:10.1101/2021.03.05.433935.
- 20 52. The 1000 Genomes Project Consortium. A map of human genome variation from
- 21 population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
- 22 53. Akhanli, S. E. & Hennig, C. Comparing clusterings and numbers of clusters by aggregation
- 23 of calibrated clustering validity indexes. *Statistics and Computing* vol. 30 1523–1544
- 24 (2020).
- 25 54. Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in
- 26 human populations. *Bioinformatics* **32**, 283–285 (2016).

- 1 55. Zhang, Y. *et al.* Local genetic correlation analysis reveals heterogeneous etiologic sharing
2 of complex traits. *Cold Spring Harbor Laboratory* 2020.05.08.084475 (2020)
3 doi:10.1101/2020.05.08.084475.
- 4 56. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering
5 clusters in large spatial databases with noise. in *Proceedings of the Second International*
6 *Conference on Knowledge Discovery and Data Mining* 226–231 (AAAI Press, 1996).
- 7 57. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection
8 for Dimension Reduction. *arXiv [stat.ML]* (2018).
- 9 58. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene
10 expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, (2018).
- 11 59. Wu, Y. *et al.* Integrative analysis of omics summary data reveals putative mechanisms
12 underlying complex traits. *Nat. Commun.* **9**, (2018).
- 13 60. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex
14 trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
- 15 61. Mancuso, N. *et al.* Probabilistic fine-mapping of transcriptome-wide association studies.
16 *Nat. Genet.* **51**, 675–682 (2019).
- 17 62. Mootha, V. K. *et al.* PGC-1 α -responsive genes involved in oxidative phosphorylation are
18 coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
- 19 63. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. *Nucleic*
20 *Acids Res.* **49**, D325–D334 (2021).
- 21 64. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more
22 comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit.
23 *Nucleic Acids Res.* **45**, W130–W137 (2017).
- 24

1 **Figures**

2

3 **Figure 1. Study Procedures and Analytic Workflow**

4 **Note:** The flow chart is divided in three columns. The first column from the left describes the analytic objectives
5 across study stages. The middle column lists the method/software used for the analysis. And the right column points
6 the reader to the Table, Figure, Supplementary Table or Supplementary Figure that reports the results of the
7 analysis.

8

9 **Figure 2. UMAP and Karyogram Plots for Meta-Locus Definitions**

10 **Note:** Panel (a), (b). UMAP plot showing 15 meta-loci and 12 meta-loci identified by Density-based scan procedures
11 across genome-wide LD independent regions for General Cognitive Ability and Non-Cognitive Skills respectively.
12 Panel (c), (d). Karyogram plots for General Cognitive Ability and Non-Cognitive Skills respectively. Color scheme in
13 each of the Karyogram plots matches those reflected in Panels (a) and (b).

14

15 **Figure 3. Violin plots for Z-score distributions of Local Genetic Correlations within each prioritized meta-locus**

16 **Note:** GlobalRgClusters: Phenotype clusters derived from partitioned clustering of global genetic correlation. Panel
17 (a). Prioritized meta-loci for General Cognitive Ability. Panel (b). Prioritized meta-loci for Non-Cognitive Skills. Panel
18 (a), (b). Bipolar: Bipolar Disorder, Anorexia: Anorexia Nervosa, Tourette's: Tourette's Syndrome, MDD: Major
19 Depressive Disorder (Howard et al., 2019), Dep-Aff: Depressive Affect, MDD_MVP: Major Depressive Disorder
20 (Million Veteran Project), Anxiety_mvmp: Anxiety Disorder (Million Veteran Project), PTSD_mvmp/pcl: Post-Traumatic
21 Stress Disorder (Million Veteran Project; Total PCL: Total PCL symptom scores)

22

23 **Figure 4. Converging gene sets identified within each General Cognitive Ability and Non-Cognitive Skills Meta-Locus**

24 **Note:** GCA: General Cognitive Ability, NCS: Non-Cognitive Skills. Panel (a), (b), Top half of circular plots are gene
25 sets, bottom half are the meta-loci. Panel (c). Overlapping gene set expanded for easy reference.

26

27 **Figure 5. BrainSpan gene expression profiles for General Cognitive Ability and Non-Cognitive Skills**

28 **Note:** GCA: General Cognitive Ability, NCS: Non-Cognitive Skills. Panel (a) scatterplot for GCA (b) scatterplot for
29 NCS. Y-axis: Normalized gene expression x-axis: Time in weeks.

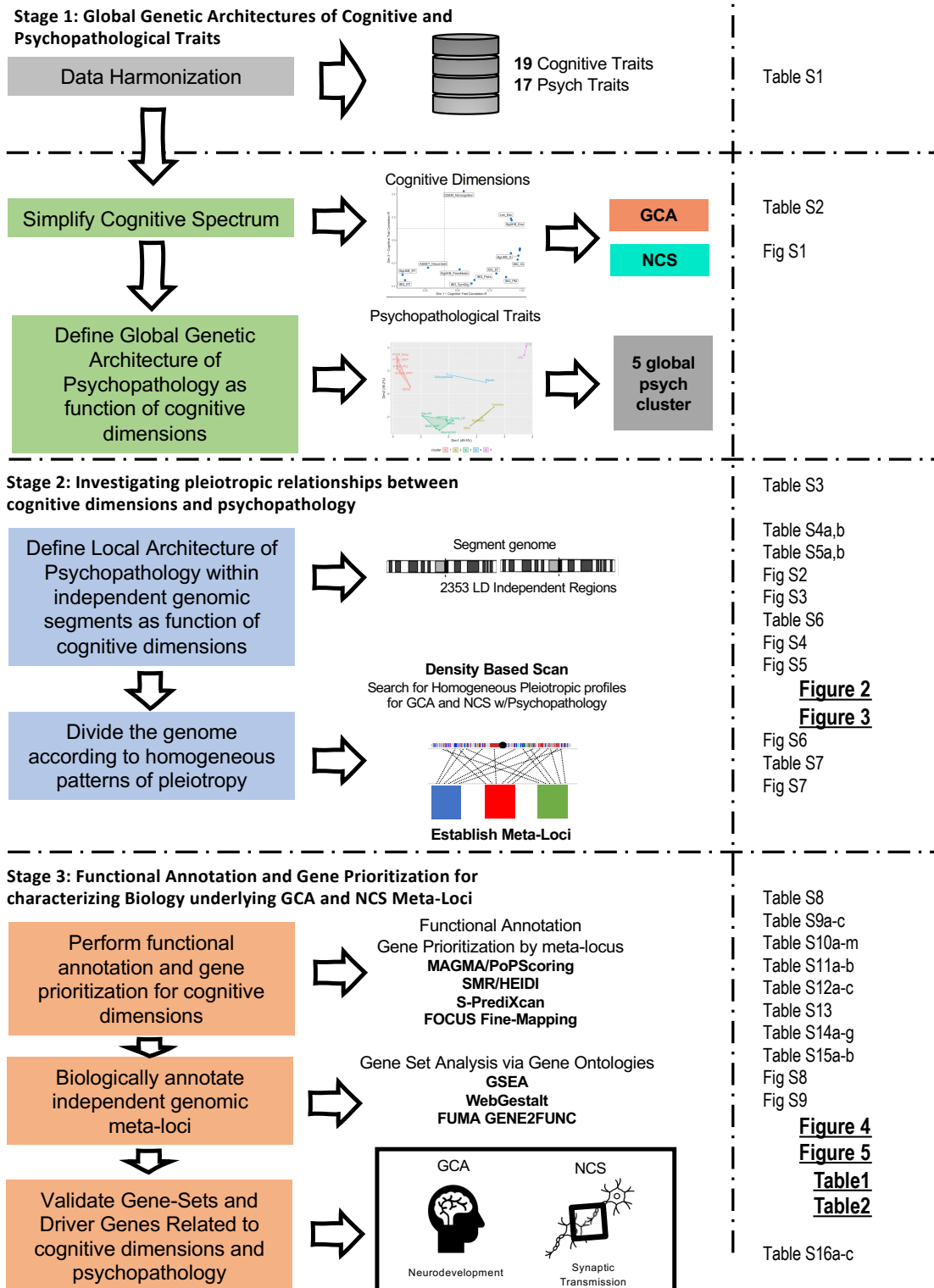
30

31

32

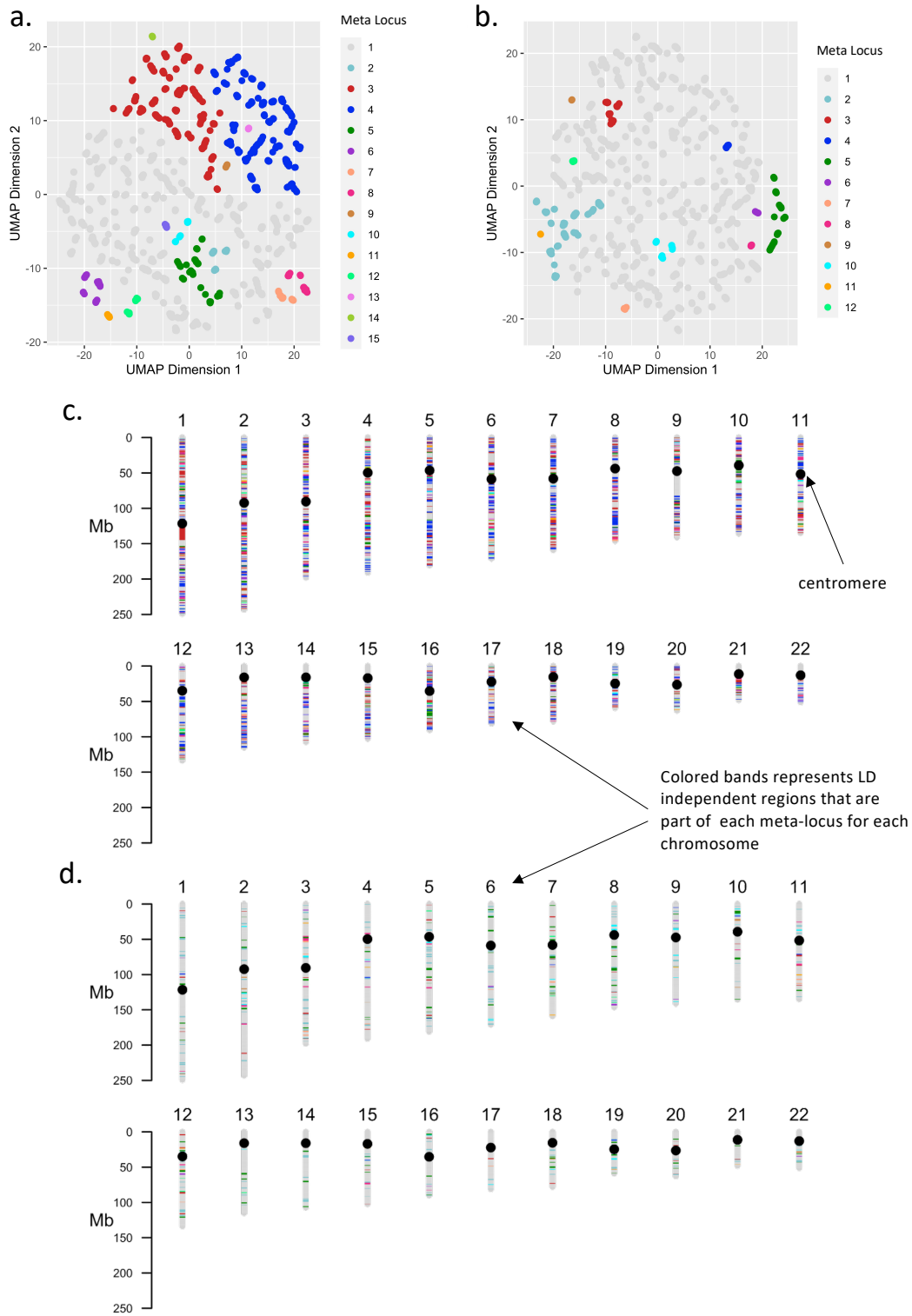
33

1 Fig-1.



2
3

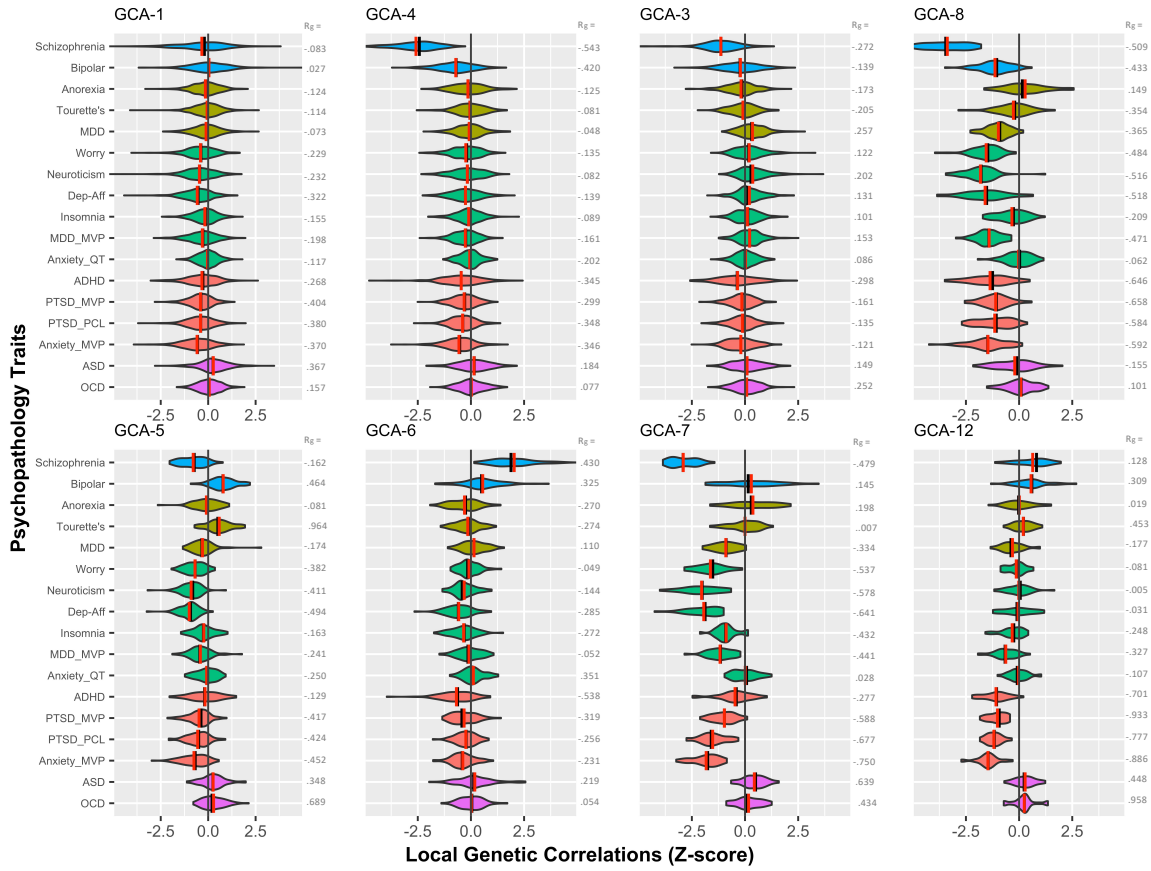
1 Fig-2.



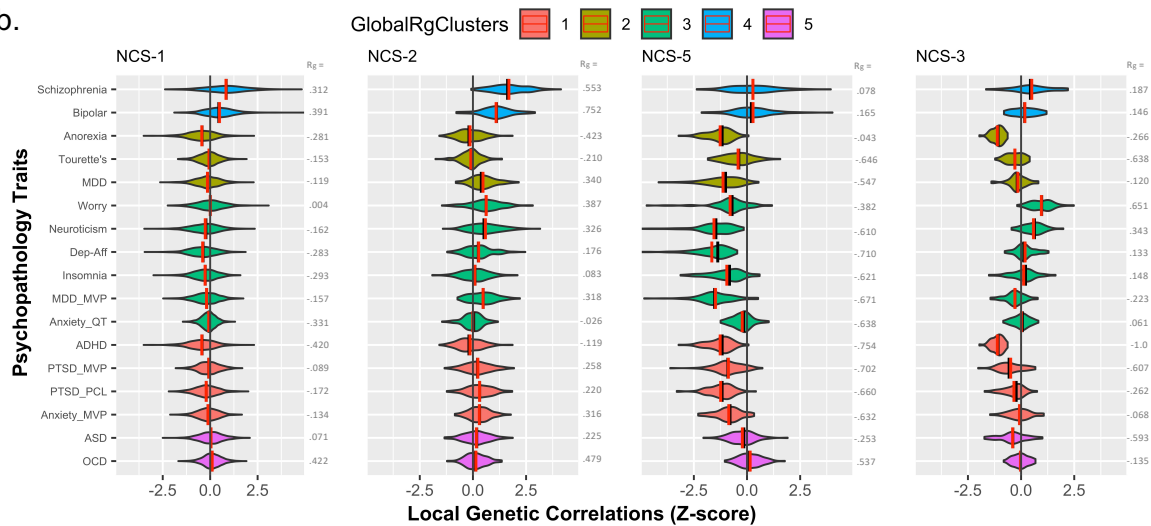
2
3

1 Fig-3.

a.

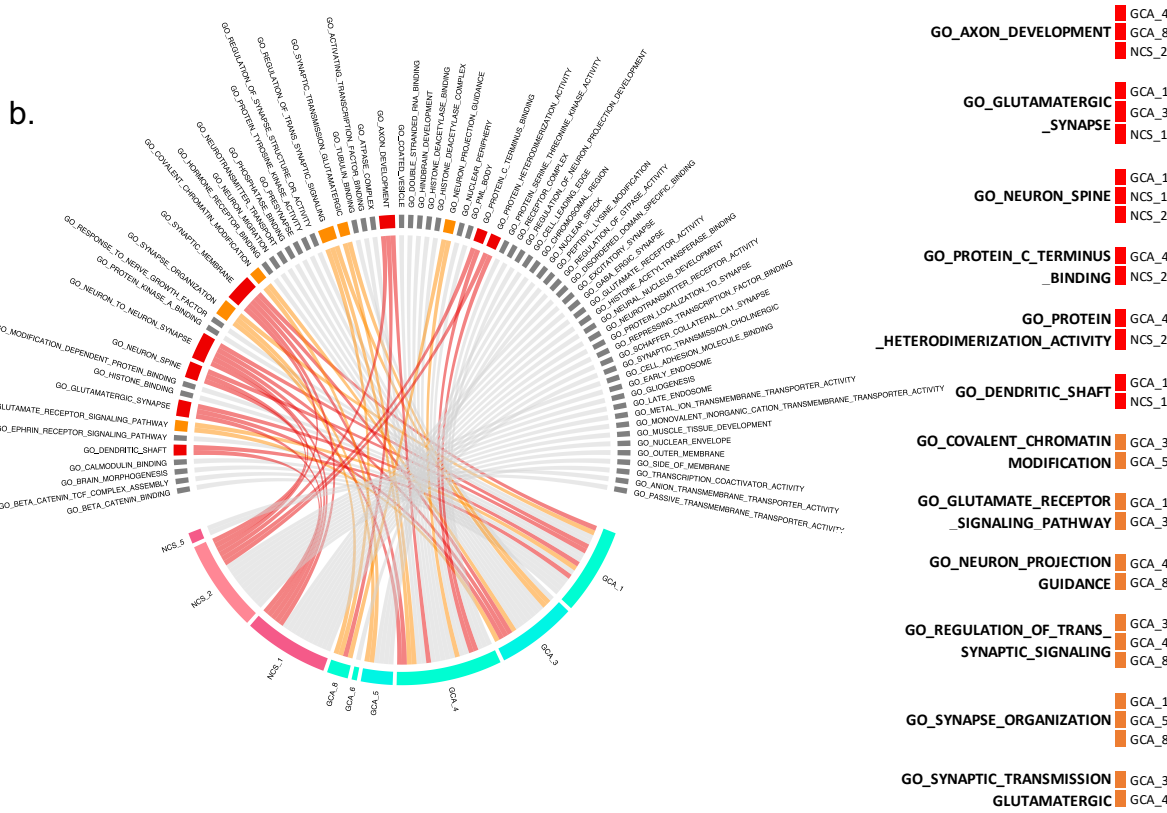
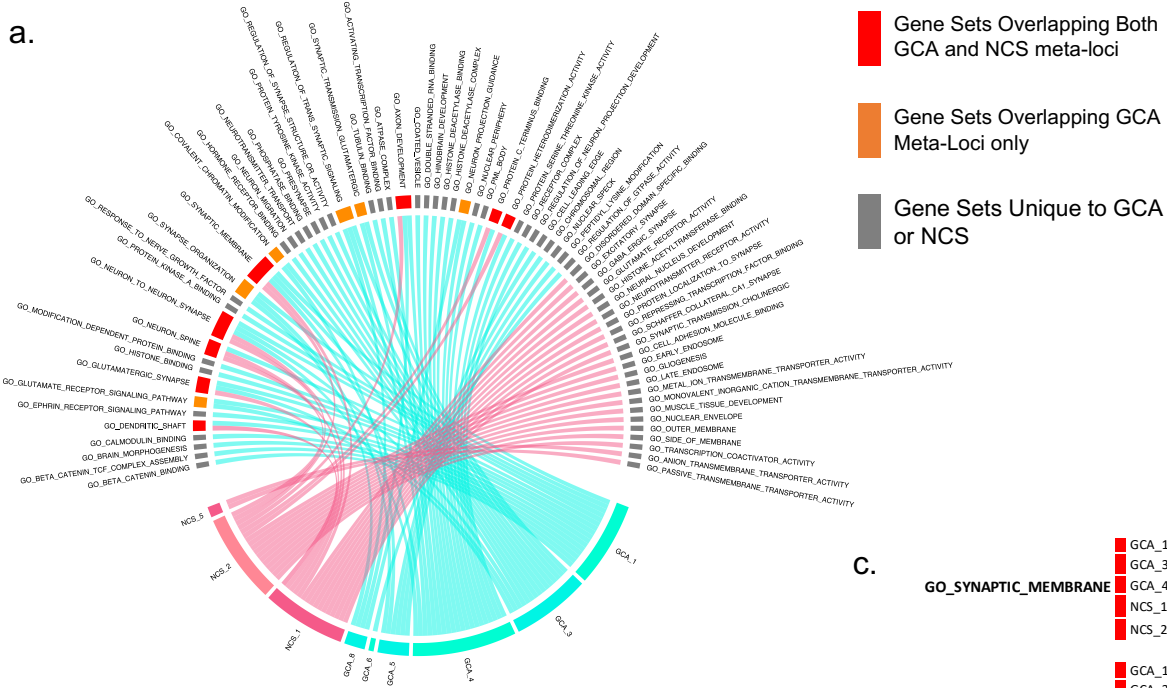


b.

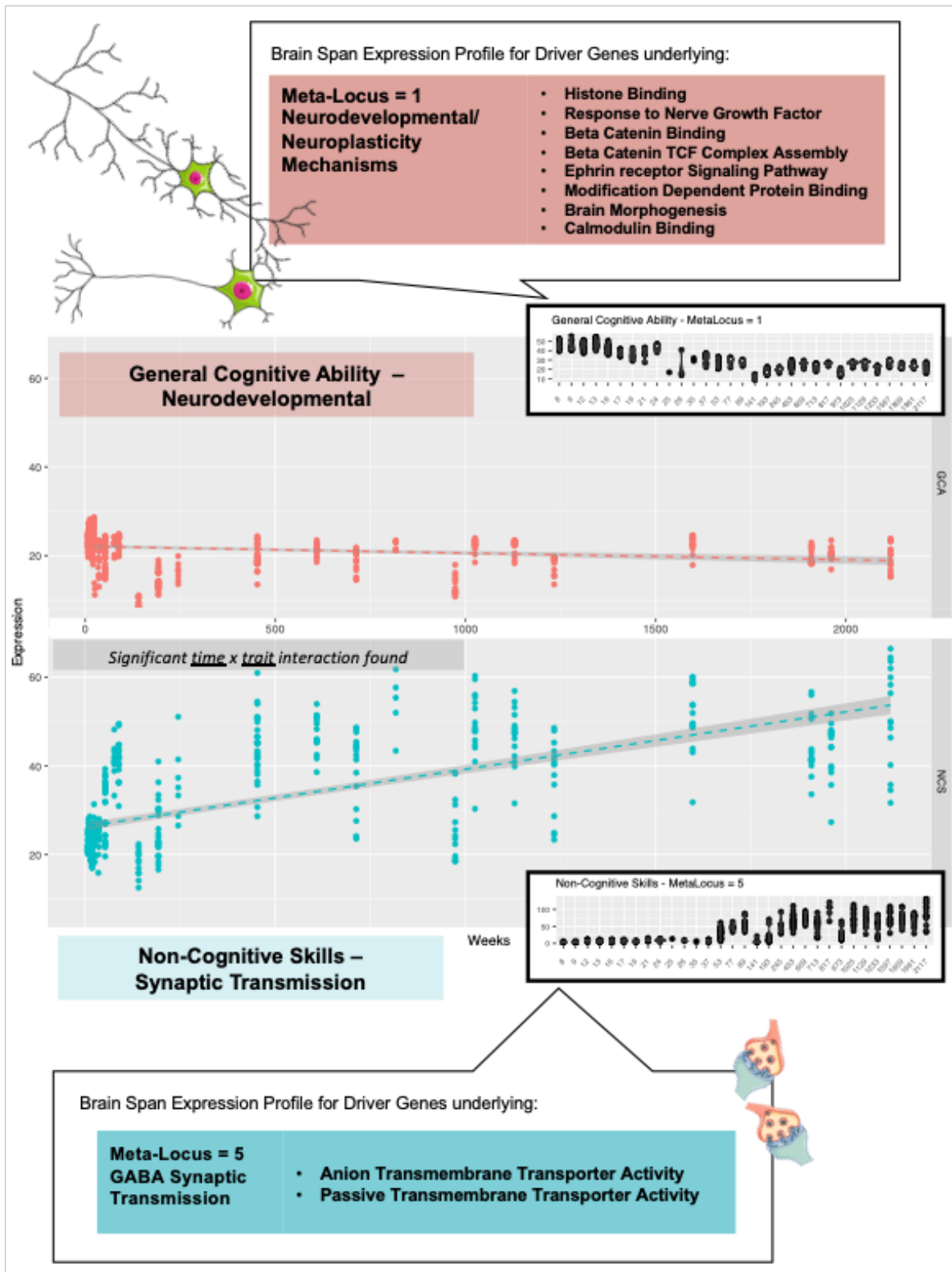


2
3

1 Fig-4.



1 Fig-5.



2
3

Table 1. General Cognitive Ability – Prioritized Meta-Loci, Gene Sets and Tier 1 Druggable Genes

GCA-1	GCA-4	GCA-3	GCA-5	GCA-8*
$h^2 = 0.104$ (45.8%)	$h^2 = 0.045$ (19.8%)	$h^2 = 0.032$ (13.9%)	$h^2 = 0.009$ (3.9%)	$h^2 = 0.010$ (4.6%)
N LD regions = 1125	N LD regions = 430	N LD regions = 377	N LD regions = 83	N LD regions = 55
Length = 1340.79 (mb)	Length = 514.32 (mb)	Length = 439.17 (mb)	Length = 101.61 (mb)	Length = 73.7 (mb)
Neurodevelopment/ Neuroplasticity	Neural Connectivity	Synaptic Structure/ Neurodegeneration	Neuronal Structure	Axonal Guidance
<ul style="list-style-type: none"> ▪ Histone Binding ▪ Response to Nerve Growth Factor ▪ Beta Catenin Binding ▪ Beta Catenin TCF Complex Assembly ▪ Ephrin Receptor Signaling Pathway ▪ Modification Dependent Protein Binding ▪ Protein Kinase A Binding ▪ Brain Morphogenesis ▪ Calmodulin Binding 	<ul style="list-style-type: none"> ▪ Coated Vesicle ▪ Double Stranded RNA Binding ▪ Activating Transcription Factor Binding ▪ Histone Deacetylase Binding ▪ ATPase Complex ▪ Nuclear Periphery ▪ Protein Serine Threonine Kinase Activity ▪ PML Body ▪ Histone Deacetylase Complex ▪ Regulation of Neuron Projection Development ▪ Receptor Complex ▪ Hindbrain development 	<ul style="list-style-type: none"> ▪ Neuron Migration ▪ Regulation of Synapse Structure or Activity ▪ Presynapse ▪ Neurotransmitter Transport ▪ Tubulin Binding ▪ Hormone Receptor Binding ▪ Protein Tyrosine Kinase Activity ▪ Phosphatase Binding 	<ul style="list-style-type: none"> ▪ Cell Leading Edge ▪ Peptidyl Lysine Modification ▪ Nuclear Speck ▪ Chromosomal Region 	<ul style="list-style-type: none"> ▪ Axon Development ▪ Neuron Projection Guidance ▪ Synapse Organization ▪ Regulation of Trans Synaptic Signaling
<p>Calcium/calmodulin dependent protein kinase: CAMK1D, CAMK2A, CAMK2G, CAMKK1,</p> <p>Ephrin Receptor: EPHA10, EPHA3, EPHA6, EPHA7, EPHB1[^],</p> <p>Glutamate Receptor: GRIN1, GRIN2B,</p> <p>Protein Kinase: PRKCB, PRKACA,</p> <p>Others: ADCY1, APH1A, CACNA1C[^], CDH2, EEF2, GSK3A, HDAC1, KMT2A, LRRK2, MAPT, MKNK2, NTRK3[^], PDPK1, PHKG2, PKN1, PPP3CA, PSEN1, RARA, ROCK1, SIRT1</p> <p><i>Total: 33 genes</i></p>	<p>Casein Kinase: CSNK1D, CSNK2A1,</p> <p>Ephrin Receptor: EPHA5, EPHA8, EPHB2[^],</p> <p>Glutamatergic Receptor: GRIK2[^], GRIN2A[^],</p> <p>Histone Deacetylase: HDAC11, HDAC3,</p> <p>Mitogen Activated Protein: MAP2K5, MAP3K10, MAP3K12, MAP3K2, MAP4K4, MAP4K5,</p> <p>Serine/Threonine Kinase STK11, STK35,</p> <p>Other: ABL1, ACVR1B, AKT2, APP, BRSK2, CDK13, CNR1, DNMT3A, DYRK1B, HIPK1, HSP90AB1, IGF1R, MARK1, NCOR2, NTRK2, PAK1, PKN2, PRKCG, RAF1, ROCK2, RPS6KB1, SHH, SLC6A4, SRPK2, TAOK3, TLK2, TNIK, VEGFA</p> <p><i>Total: 45 genes</i></p>	<p>CDC-like Kinase: CLK2, CLK4,</p> <p>Ephrin Receptor: EPHA4, EPHB2[^], EPHB6</p> <p>Glutamatergic Receptor: GRIA1, GRIK2[^], GRIN2A[^],</p> <p>Solute Carrier family: SLC6A1, SLC9A1,</p> <p>Others: AKT1, CAMK2B, CDK5, CSK, ERBB4[^], ESR1, FGFR3, FKBP4, HTT, JAK2, MAPK14, MET[^], NTRK3[^], RXRA, SRC, TP53</p> <p><i>Total: 26 genes</i></p>	<p>EPOR</p> <p><i>Total: 1 gene</i></p>	<p>PTK2, GRIK3, PIK3CA, NISCH, EPHA10, NCAM1, VAMP2, DRD2, CACNA1D</p> <p><i>Total: 9 genes</i></p>

Table 2. Non-Cognitive Skills – Prioritized Meta-Loci, Gene Sets and Tier 1 Druggable Genes

NCS-1	NCS-2	NCS-5
$h^2 = 0.244$ (78.3%)	$h^2 = 0.024$ (7.6%)	$h^2 = 0.020$ (6.3%)
N LD regions = 1926	N LD regions = 162	N LD regions = 89
Length = 2290.00 (mb)	Length = 191.66 (mb)	Length = 116.4 (mb)
Synaptic Activity/ Neurotransmission	Synaptic Transmission/ Action Potentiation	GABA-ergic Synaptic Transmission
<ul style="list-style-type: none"> • Neurotransmitter Receptor Activity • Disordered Domain Specific Binding • Schaffer Collateral CA1 Synapse • Repressing Transcription Factor Binding • Excitatory Synapse • Protein Localization to Synapse • Gaba-ergic Synapse • Neural Nucleus Development • Histone Acetyltransferase Binding • Synaptic Transmission Cholinergic • Glutamate Receptor Activity 	<ul style="list-style-type: none"> • Metal Ion Transmembrane Transporter Activity • Monovalent Inorganic Cation Transmembrane Transporter Activity • Muscle Tissue Development • Early Endosome • Late Endosome • Outer Membrane • Side of Membrane • Cell Adhesion Molecule Binding • Nuclear Envelope • Transcription Coactivator Activity • Gliogenesis 	<ul style="list-style-type: none"> • Anion Transmembrane Transporter Activity • Passive Transmembrane Transporter Activity
<p>Cyclin Dependent Kinase: CDK5, CDK5R1, Cholinergic Receptor: CHRM1, CHRM2, CHRNA1, CHRNA2, CHRNA3, CHRNA4, CHRNA7, CHRNB1, CHRNB2, CHRNB4, CHRNE, CHRNG, Dopamine Receptor: DRD1, DRD5, GABAergic Receptor: GABRA5, GABRB2[^], GABRB3, GABRG3, Glutamatergic Receptor: GRIA1, GRIA2, GRIA4, GRIK2, GRIN1, GRIN2C, GRIN2D, GRM5, Histone Deacetylase: HDAC1, HDAC2, HDAC4, HDAC5, HDAC7, HDAC9[^], Heat Shock Protein: HSP90AA1, HSP90AB1, 5-hydroxytryptamine receptor: HTR1B, HTR1F, Solute Carrier Family: SLC6A1, SLC6A9, Others: ADORA1, BCL2[^], BCR, CNR1, EPHA7, ERBB4[^], FYN, HIF1A, HRAS, MAPT, MDM2, NQO1, PDE4B[^], PPP3CA, PTK2B, SV2A, TP53 <i>Total: 57 genes</i></p>	<p>Apoptosis regulator: BCL2[^], BCL2L1, Ephrin receptor: EPHA2, EPHA4, EPHB1, Intercellular Adhesion Molecule: ICAM1, ICAM3, Other: ATP1A2, GHR[^], HMGCR, IL11RA, ITGAL, KCNK3, MAPK3, PAK6, PPARG, PTPRC, RAF1, SIGMAR1 <i>Total: 19 genes</i></p>	<p>GABAergic Receptor: GABRA1, GABRA2, GABRA4, GABRA6, GABRB1, GABRB2[^], GABRG1, GABRG2, GRIN2B, Other: KCNB2, P2RX4, SLC12A5, SNAP25 <i>Total: 13 genes</i></p>